



Classification of Alzheimer's disease from MRI Images using CNN based Pre-trained VGG-19 Model

Manimurugan S

¹Department of Computer Engineering, Faculty of Computers and Information Technology, University of Tabuk, Saudi Arabia.

***Corresponding Author: semanimurugan@gmail.com, mmurugan@ut.edu.sa**

Received: 15.06.2020,
Revised: 21.07.2020,
Accepted: 19.08.2020,
Published: 31.09.2020

DOI:
[10.53409/mnaa.jcsit20201205](https://doi.org/10.53409/mnaa.jcsit20201205)

Abstract: The most typical form of dementia is Alzheimer's disease (AD), which can permanently cause memory cell damage. Since Alzheimer's is a progressive disease, various problems increase over time. For this reason, diagnosing the disease early is very necessary. The primary objective of this work is to use Magnetic Resonance Imaging to diagnose Alzheimer's disease. A Convolution Neural Network (CNN) model, VGG-19, is proposed in this research. VGG-19 is typically a pre-trained deep transfer learning model that is used here to use MRI images to identify Alzheimer's disease. The primary processes performed in this research are preprocessing, feature extraction, and classification. From the OASIS database, the dataset used in this work is obtained. A total of 373 MRI images are used for evaluation. For training, 80% of data and testing, 20% of data are used in this model for performance analysis. The proposed VGG-19 model is evaluated and compared with existing deep learning-based other CNN models like AlexNet, GoogLeNet, and VGG-16. The performance metrics like accuracy, recall, precision, specificity, and f-measure are evaluated to estimate the performance analysis of the model. For every validation, both training and testing results are evaluated and compared. The model has obtained 96.04% training accuracy and 95.82% testing accuracy, which was 1.9% to 5.3% higher than the AlexNet and GoogLeNet models in performance. However, this model is proposed to classify the AD from brain MRI images and obtain better validation results.

Keywords: Alzheimer's disease, Dementia, CNN, VGG-19, OASIS database, MRI.

I. INTRODUCTION

Dementia is a disorder in which memory, actions, perception, and the ability to perform daily tasks deteriorate. The dementia affects the older people mostly and it is not a general part of aging. The most common type of dementia is AD and might lead to 60-70% of cases. Around 50 million people have dementia worldwide, according to the WHO, with about 60 percent living in low- and middle-income countries. There are about 10 million new cases each year. The average percentage at a given time of the general population aged 60 and over with dementia is between 5-8 percent. It is predicted that the total number of people with dementia will surpass 82 million in 2030 and 152 million in 2050. The loss of tissues and death of nerve cells in the brain is caused by AD, leads to human memory loss and having a bad effect on

the output of everyday life activities like writing, communicating, and reading. Patients with AD can also have trouble in recognizing their members of family. Patients in the mild cognitive stage act aggressive, but people in the final AD state suffer from cardiac arrest and death-related respiratory dysfunction [1]. For dementia scaling, the GDS (global deterioration scale) was widely utilized. This scale additionally divides dementia into 7 stages, considering stages 1-3 as stages of pre-dementia, and considering stages 4-7 as dementia.

1. No cognitive decline
2. Age-Associated Memory Impairment
3. Mild Cognitive Impairment
4. Mild Dementia
5. Moderate Dementia
6. Moderately Severe Dementia
7. Severe Dementia

In the field of detecting AD, neuroimaging techniques help the evaluation of brain changes and are promising. MRI plays a vital role in neuroimaging techniques in offering both the anatomical and physiological functions of the body. In diagnosing AD, MRI is the most frequently used technique where any defects present in the brain can be detected [1]. MRI-based images provide multi-modal information that is appropriate for clinical purposes of the brain's function and structure.

Many Computer-Aided Diagnostic Systems (CADs) have been developed by researchers to

precisely detect and classify the extracted features of AD. Otherwise, additional time and effort by specialists is needed to process the extracted features. Researchers have recently developed deep learning techniques/models to explicitly extricate features from medical images. Deep learning models have obtained significant performance in medical images, like X-ray, MRI, CT, microscopy, and mammography. The key emphasis of these models or approaches was on binary classification, which indicates whether or not the patient is suffering from AD [2].

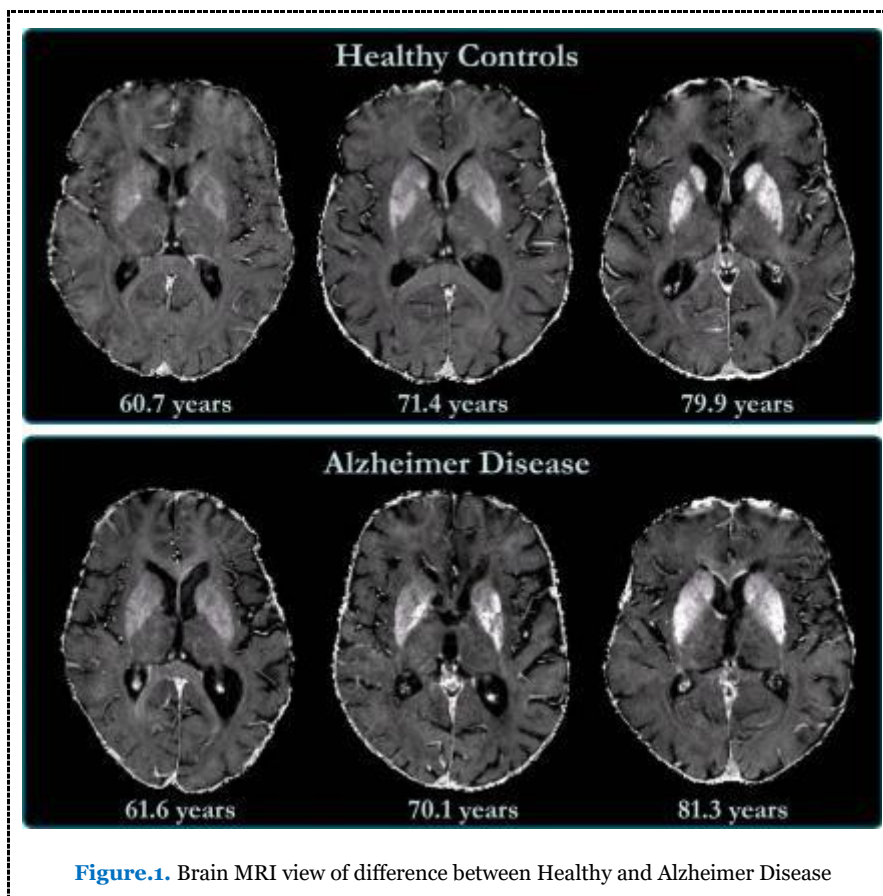


Figure.1. Brain MRI view of difference between Healthy and Alzheimer Disease

In this research, a pre-trained deep convolution neural network model VGG-19 with transfer learning is proposed to analyze and classify AD. This research's main objective is to detect AD based on a CNN model VGG-19 using MRI images. The MRI images collected from the OASIS database are used to evaluate the proposed model in detecting AD. The assessment of performance analysis is based on accuracy, precision, recall, specificity, and F-measure.

II. RELATED WORK

Braulio Solano Rojas et al. proposed a three-dimensional Densenet-121 architecture for AD

identification using MRI images. The images were gathered and preprocessed from the ADNI database. Some good performance results have been obtained by this prediction model, and still, to produce more than average results, the model can be improved. A downside may be the reduction of the dataset used for training and testing [3].

Atif Mehmood et al. developed the Siamese-CNN (SCNN) for AD classification. The model was totally similar but joins two parallel layers of modified VGG16, called "Siamese." In this work, the OASIS dataset was used. The initial stage was preprocessing of data and augmentation, the next was extracting

features, and finally, classification. To obtain the maximum features on a limited dataset, the VGG16 and an additional Conv layer were added to the model. A CNN based VGG-16 model was used for classification. Four dementia stage categories have been classified and better performance has been achieved [4].

Taranjit Kaur and Tapan Kumar Gandhi used a pre-trained model DCNN-VGG-16 with transfer learning (TL) for the identification of abnormalities in brain MRI. The features extricated by CNN were highly dependent on the training dataset's size. From pre-trained models, weights of the 'conv' layer were utilized in TL, and just the final layers were trained with data from the new classes. In the context of the elimination of the ROI delineation's preprocessing stages, feature extraction, and selection, the performances were better to the current conventional methods recorded for the classification of brain image [5].

Xiaoling Lu et al. used the MobileNet, a CNN technique for the classification of AD using MRI images. The MRI images were initially preprocessed and using the convolution layer portion, the bottleneck features were extracted. Then to create a new network structure, the newly built completely connected layer was connected. TL was used to load weights on new networks and train networks for pre-training. The MRI images were eventually identified and compared with the model VGG-16. In this work, MobileNet is observed to be less complex and better than VGG-16 [6].

Muhammed Yildirim and Ahmet Cinar proposed a hybrid classification model based on CNN for classifying AD using MRI images. For this hybrid model, the ResNet50 model was used as a basis, and the output was individually tested using different CNN models such as AlexNet, ResNet50, DenseNet201, and VGG16. The last five layers of Resnet50 were removed in the hybrid model, ten new layers were added in place of those layers removed, and the number of layers increased from 177 to 182. The efficiency of the hybrid model has achieved greater accuracy than others at 90 percent compared to all models [7].

Ronghui Ju et al. presented a model for early recognition of AD using deep learning with brain networks. The brain network was built by evaluating the brain region's functional connectivity using data from resting-state functional MRI (R-fMRI). To detect an early AD state, normal aging from mild cognitive impairment, a targeted autoencoder network was implemented. The proposed technique

effectively revealed discriminative brain network features and produced a steady AD detection classifier [8].

Marcia Hon and Naimul Mefraz Khan proposed the TL-based AD detection technique using VGG-16 and Inception v4 models from MRI images. The two models were both trained and evaluated and used the entropy-based approach to select the training dataset. In detecting AD, the inception model outperformed the VGG-16 model with an accuracy of 96.25% [9].

Xin Hong et al., using the LSTM technique, proposed a model for predicting AD. A special type of recurrent neural network that could connect previous data to the current task is long short-term memory. The LSTM network with completely linked layers and activation layers was therefore constructed in order to encode the temporal relationship among features and the next stage of AD [10]. F.J. Martinez Murcia et al. proposed a new research data analysis of AD using the deep convolutional autoencoder technique. Using regression and classification analysis, the distribution of the extracted features in various combinations were analyzed and visualized. The impact on the brain of each co-ordinate of the autoencoder manifold was evaluated. The assessment was performed by using the ADNI MRI dataset and achieved an accuracy of over 80 percent [11].

Mingxia Liu et al. proposed a weakly-supervised densely linked neural network (wiseDNN) for brain disease prognosis using baseline MRI data and incomplete clinical scores [12]. Naimul Mefraz Khan et al. proposed a TL-based technique for AD diagnosis from MRI. With layer-wise tuning, the network was fine-tuned, where only a group of pre-defined layers was trained on MRI images. The entropy approach was utilized to decrease the training dataset size and achieved 95.19 percent accuracy [13].

III. PROPOSED METHODOLOGY

CNN is extensively recognized for its capability to execute highly accurate medical image classification in deep learning. However, compared to traditional machine learning techniques, the most significant benefit of CNNs is that it does not need manual extraction of features because CNNs can automatically extract the features and then classifies the AD stages. The principle of TL is to train a pre-trained CNN to use a smaller dataset from a different problem to learn new image representations. TL is a framework for machine learning design; it is not a model or

technique for machine learning. For pre-trained models, this design approach usually applies. These pre-trained models are based on Deep Evolution Neural Networks. In deep learning, this technique involves the initial training of a CNN using large-scale training datasets for a classification issue. Figure 2 demonstrates the architecture of the proposed VGG-19 model.

As shown in Figure 3, VGG-CNN has six main structures, mostly comprised of multi-connected convolutional layers and fully-connected layers. 3*3 is the convolutional kernel dimension, and 224*224*3 is the input size. In general, every layer is concentrated at

16~19. As a preprocessing model, VGG-19 was used. It has been increased in network complexity compared with conventional CNNs. It uses an alternating structure, which is well over a single convolution, with several convolution and non-linear activation layers. The layer structure can extract better features of image, using Maxpooling for downsampling, and adjusting the ReLU as the activation function, selecting the largest value in the image region as the region's pooled value. The downsampling layer is primarily utilized to enhance the image anti-distortion capability of the network while maintaining the main sample features and minimizing the number of parameters [14].

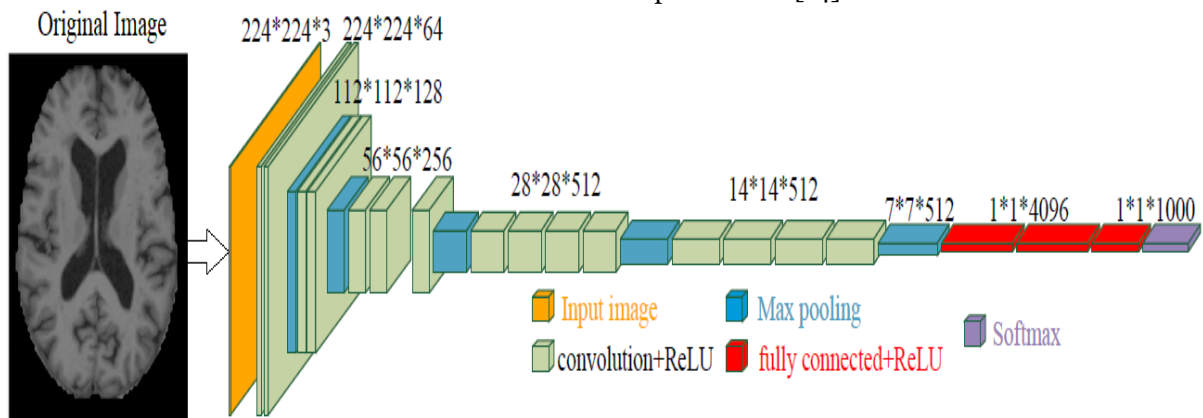


Figure.2. Proposed VGG-19 Architecture Model

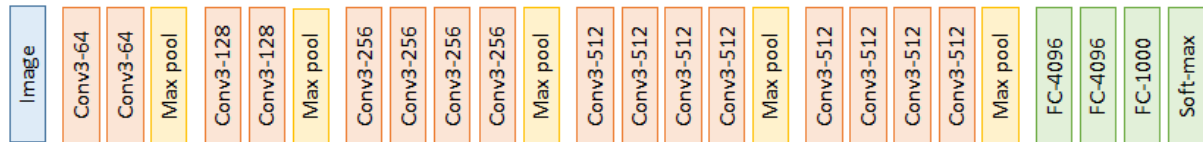


Figure.3. Layers of VGG-19

For the downsampling layer, the expression is presented in equation (1). Among them, $down(X_j^{(n-1)})$ is the maximum pooling sampling function. τ_j^n is the coefficient according to the j th feature map of the n th layer, and $f(\tau_j^n down(\tau_j^n down(X_j^{(n-1)})) + b_j^{(n)})$ is the ReLU activation function.

$$X_{pj}^{(n)} = f(\tau_j^n down(X_j^{(n-1)}) + b_j^{(n)}) \quad (1)$$

The activation function of ReLU is expressed as follows,

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (2)$$

The activation function of the softmax layer is expressed as follows,

$$f(x_j) = \frac{e^{x_j}}{\sum e^{x_i}} \quad (3)$$

In these above equations, $f(x)$ is the activation function, and x is the activation function input.

The initial two layers are convolutional layers with 3x3 filters, and 64 filters are used in the first two layers, so it finishes with a volume of 224x224x64 since the same convolutions have been used (height and width are the same). So this (CONV64)x2 reflects that there are two Conv layers with 64 filters in the model.

- The filters are always 3x3 with a stride of 1, and the same convolutions are always applied.
- A pooling layer is then used, which lowers a volume's height and width: it goes from 224x224x64 down to 112x112x64.
- Then it uses a few more layers of Conv. 128 filters are used here, and the same convolutions are used; 112x112x128 would be a new dimension.
- Then, a pooling layer is added so that the new dimension will be 56x56x128.
- Two conv layers with 256 filters are added next.

- A pooling layer
- A few more conv layers with 512 filters
- A pooling layer
- A few more conv layers with 512 filters
- A pooling layer
- In the end, the model has its final $7 \times 7 \times 512$ in a Fully-connected layer (FC) with 4096 units and one of 1000 classes in a softmax output.

With a larger collection of data over a smaller one, CNN typically performs well. In those CNN applications where the dataset is not large transfer learning (TL) may be useful. TL's concept utilizes the learned model from large datasets such as ImageNet for applications with comparatively smaller datasets. This removes the need for a large dataset and decreases the lengthy training time as needed when generated from scratch by the deep learning algorithm. TL is a deep learning technique that uses a model trained as a starting point for a single assignment to train a model for a similar assignment. Fine-tuning a network with TL is usually much faster and easier than training a network from scratch. It enables training models using similar small labeled data by leveraging standard models that have already been trained on large datasets. It is possible to dramatically decrease the training time and computing resources. With TL, the model does not need to be trained for as many epochs (a complete training period on the entire dataset) as a new model.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed CNN's performance analysis with the VGG-19 model is assessed using the dataset in this section. The model is evaluated using parameters such as accuracy, recall, precision, specificity, and f-measure. Also, a comparative analysis is conducted for the validation of the model proposed. The output is compared to other current deep learning models used for classification, such as AlexNet, GoogLeNet, and VGG-16. On the MATLAB 2019a Simulink toolbox, all the experiments are simulated and analyzed. The dataset is split into 80 percent for training and 20 percent for testing.

4.1. Dataset Description

The proposed performance review of CNN with the VGG-19 model is assessed in this section using the dataset. Using parameters such as accuracy, recall, precision, specificity, and f-measure, the model is evaluated. A comparative analysis is also performed for the validation of the proposed model. The

performance is compared with other existing deep learning models, such as AlexNet, GoogLeNet, and VGG-16, used for classification. All the experiments are simulated and analyzed on the MATLAB 2019a Simulink toolbox. The dataset is split into 80 percent for training and 20 percent for testing.

In this research, the OASIS dataset is used for the evaluation. The Open Access Series of Imaging Studies (OASIS) aims to make neuroimaging data sets of the brain freely available to the scientific community. The dataset consists of 150 subjects aged 60 to 96 years in a longitudinal collection. For a total of 373 imaging sessions, every subject was diagnosed for two or more visits, separated by at least one year. 3 or 4 individual T1-weighted MRI scans acquired in single scan sessions are included for each subject. Across the study, 72 of the subjects were labelled as non-demented. At the time of their initial visits, 64 of the subjects included were labelled as demented and remained so for subsequent scans, including 51 people with mild to moderate AD. Based on the CNN model, data preprocessing was the significant part of obtaining effective and precise results for those algorithms. The image size of the OASIS dataset was 256×256 , but an image size of 224×224 is needed for the proposed CNN model.

4.2. Performance Metrics

In this research, the proposed CNN with the VGG-19 model is proposed and compared with other deep learning CNN models such as AlexNet, GoogLeNet, and VGG-16. To estimate the performance analysis of the model, the performance metrics such as precision, accuracy, recall, specificity, and f-measure are evaluated. Both training and testing outcomes are measured and compared with every validation. The primary objective of this research is to identify AD from MRI brain images, which can be helpful in assessing the patient's abnormality. This model's outcome may be dependent on the result detected as normal or abnormal. True positive (TP), true negative (TN), false positive (FP), and false negative (FN) are correctly evaluated to predict this model's outcome.

TP: It refers the total correct predictions in abnormal cases.

FP: It refers the total incorrect predictions in abnormal cases.

TN: It refers the total correct predictions in normal cases.

FN: It refers the total incorrect predictions in normal cases.

Accuracy is the estimate of the performance subset by the model. It is the primary performance metric used to measure the classification process efficiency. Where both the positive and negative groups are equally significant, it is generally used for evaluation. It is calculated using the equation below,

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision is a predictive value that is positive. The preciseness of the classification model is computed by using it. It is the estimate of the correctly predicted positive observation's cumulative predictive positive value. The lower precision value reflects a large number of false positives, which affects the classification model. The measure of precision can be computed using the equation below,

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Sensitivity is also known as recall. It is the proportion of correctly predicted positive observation of the overall positive predictive value. The lower recall value reflects a large number of false-negative values, which affects the classification model. The recall estimation can be calculated using the equation below,

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

According to this model, specificity is the prediction that healthy subjects do not have an abnormality. It is the percentage of subjects with no injury/trauma that is tested as abnormal. The specificity estimation can be calculated using the equation below,

$$Specificity = \frac{TN}{TN+FP} \quad (7)$$

The F-measure estimates the performance accuracy and is specified as the weighted harmonic mean of the precision and the recall. The accuracy does not take into account how the data is distributed. The f-measure is then used to manage the distribution problem with accuracy. The f-measure estimation can be calculated using the following equation,

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

Table.1. Performance Analysis of Training Evaluation

Model	Accuracy	Precision	Recall	Specificity	F-measure
AlexNet	94.13	93.81	93.64	95.25	93.72
GoogLeNet	95.66	94.22	93.80	95.90	95.39
VGG-16	95.84	94.90	94.92	96.02	95.56

VGG-19 96.04 95.96 95.75 96.18 95.81

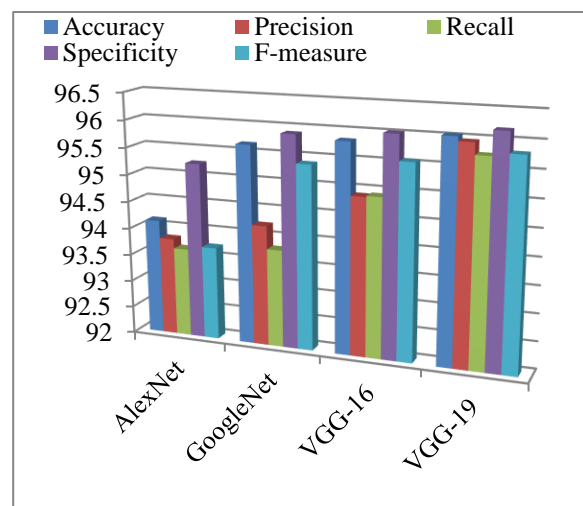


Figure.4. Performance Analysis of Training Data

Table.2. Performance Analysis of Testing Evaluation

Model	Accuracy	Precision	Recall	Specificity	F-measure
AlexNet	89.45	88.74	88.23	90.07	89.19
GoogLeNet	92.90	91.55	91.06	93.12	91.09
VGG-16	94.91	93.32	93.35	95.31	94.24
VGG-19	94.82	93.24	93.13	95.14	94.10

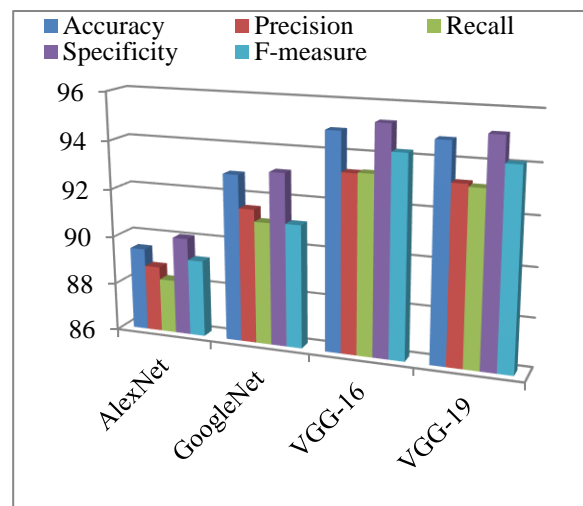


Figure.5. Performance Analysis of Testing Data

As shown in table.1 and 2, the proposed VGG-19 model achieved better performance in training and testing for classifying the ACL tear MRI images. The model obtained 96.04% training accuracy and 95.82% testing accuracy, which is 1.9% to 5.3% increased performance than the other existing compared models. The AlexNet, GoogLeNet, and VGG-16 achieved 94.13%, 95.66%, 95.84% in training and 89.45%, 92.90%, 94.91% in testing

performance respectively. The VGG-16 and VGG-19 model achieved better performance in testing compared to the performance. The performance was the same as accuracy for other parametric evaluations like precision, recall, specificity, and f-measure.

Compared with both training and testing results, the AlexNet model achieved the least performance in every parameter. The VGG-16 model has achieved close results to the proposed model and better accuracy on testing performance. The graphical chart of the comparison is plotted in figure.4 and 5.

V. CONCLUSION AND FUTURE WORK

In this research, a CNN based VGG-19 model was proposed for classifying the AD using MRI images. The proposed model was executed into four stages; Data preprocessing is the initial stage. After preprocessing, the next stage is to extract the image's features and then deliver the extracted features into the classifier to train. The model after training is finally tested. The dataset used in this work was collected from the OASIS database. It is an open-access database for the brain MRI dataset. A total of 373 MRI images were used for evaluation. For training, 80% of data (298 images) and testing, 20% of data (75 images) were used in this model for performance analysis. The proposed VGG-19 model was evaluated and compared with existing deep learning-based other CNN models like AlexNet, GoogLeNet, and VGG-16. The performance metrics like accuracy, precision, recall, specificity, and F-measure were evaluated to estimate the performance analysis of the model. For every validation, both training and testing results were evaluated and compared. The model has obtained 96.04% training accuracy and 95.82% testing accuracy, which was 1.9% to 5.3% higher than the AlexNet and GoogLeNet models in performance. However, this model was proposed to classify the AD from brain MRI images and obtained better validation results. In future, the proposed model can be used to find any abnormalities present in various body organs by using different datasets like a brain tumor, heart disease, lung cancer, etc. The performance can be further increased by implementing a new deep transfer learning model with better classification performance.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

None.

REFERENCES

- [1] Mamata L and Rashmi P. (2018). A Survey on Classification Methods of Brain MRI for Alzheimer's disease. *Int J Eng Res Technol*. Vol.7, No.5, pp.339-348.
- [2] Suhad Al-Shoukry, Taha H. R, and Nasrin M. M. (2020). Alzheimer's Diseases Detection using Deep Learning Algorithms: A Mini-Review. *IEEE Access*. Vol.8, pp.77131-77141, DOI: 10.1109/ACCESS.2020.2989396.
- [3] Solano-R B., Villalón-F R., Marín-R G. (2020). Alzheimer's Disease Early Detection Using a Low-Cost Three-Dimensional Densenet-121 Architecture. In: Jmaiel M., Mokhtari M., Abdulrazak B., Aloulou H., Kallel S. (eds) *The Impact of Digital Technologies on Public Health in Developed and Developing Countries*. ICOST 2020. *Lecture Notes in Computer Science*, Vol.12157, pp.3-15. Springer, Cham. <https://doi.org/10.1007/978-3-030-51517-11>.
- [4] Atif M, Muazzam M, Muzaffar B, and Yang S. (2020). A Deep Siamese Convolution Neural Network for Multi-Class Classification of Alzheimer Disease. *Brain Sciences*. Vol.10, No.84, pp.1-15.
- [5] Taranjit K and Tapan K G. (2019). Automated brain image classification based on VGG-16 and transfer learning. 2019 International Conference on Information Technology (ICIT), pp.94-98, DOI: 10.1109/ICIT48102.2019.00023.
- [6] Xiaoling L, Haifeng W, and Yu Z. (2019). Classification of Alzheimer's disease in MobileNet. *Journal of Physics: Conference Series*. Vol.1345, 042012, DOI:10.1088/1742-6596/1345/4/042012.
- [7] Muhammed Y and Ahmet C. (2020). Classification of Alzheimer's disease MRI Images with CNN Based Hybrid Method. *Ingénierie des Systèmes d'Information*. Vol.25, No.4, pp.413-418. <https://doi.org/10.18280/isi.250402>.
- [8] Ronghui J, Chenhui H, Pan Z, and Quanzheng L. (2019). Early Diagnosis of Alzheimer's disease based on Resting-State Brain Networks and Deep

- Learning. IEEE/ACM T Comput Biol Bioinform, Vol.16, No.1, pp.244-257, DOI: 10.1109/TCBB.2017.2776910.
- [9] Marcia H and Naimul M K. (2017). Towards Alzheimer's disease Classification through Transfer Learning. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp.1166-1169. DOI: 10.1109/BIBM.2017.8217822.
- [10] Xin H, Rongjie L, Chenhui Y, Nianyin Z, Chunting C, Jin G, and Jane Y. (2019). Predicting Alzheimer's Disease Using LSTM, IEEE Access, Vol.7, pp.80893-80901, DOI: 10.1109/ACCESS.2019.2919385.
- [11] F.J. Martinez-Murcia, A. Ortiz, J.M. Gorriz, J. Ramirez, and D. Castillo-Barnes. (2020). Studying the Manifold Structure of Alzheimer's disease: A Deep Learning Approach Using Convolutional Autoencoders, IEEE J Biomed Health, Vol.24, No.1, pp.17-26, DOI: 10.1109/JBHL.2019.2914970.
- [12] Mingxia L, Jun Z, Chunfeng L, and Dinggang S. (2020). Weakly Supervised Deep Learning for Brain Disease Prognosis Using MRI and Incomplete Clinical Scores. IEEE T Cybernetics. Vol.50, No.7, pp.3381-3392, DOI: 10.1109/TCYB.2019.2904186.
- [13] Naimul M K, Nabila A, and Marcia H. (2019). Transfer Learning with Intelligent Training Data Selection for Prediction of Alzheimer's disease. IEEE Access. Vol.7, pp.72726-72735, DOI: 10.1109/ACCESS.2019.2920448.
- [14] Jian X, Jia W, Shaozhong C, and Bilong L. (2020). Application of a Novel and Improved VGG-19 Network in the Detection of Workers Wearing Masks, Journal of Physics: Conference Series Vol.1518, 012041, DOI: 10.1088/1742-6596/1518/1/012041.

Cite this article as: Manimurugan S., Classification of Alzheimer's disease from MRI Images using CNN based Pre-trained VGG-19 Model. J. Comput. Sci. Intell. Technol. 2020; 1(2): 34-41. ©JCSIT, MNAA PUB WORLD, 2020.